Sleep Medicine Reviews 28 (2016) 55–68

Contents lists available at ScienceDirect

Sleep Medicine Reviews

journal homepage: www.elsevier.com/locate/smrv

CLINICAL REVIEW

The effect of treating obstructive sleep apnea with positive airway pressure on depression and other subjective symptoms: A systematic review and meta-analysis

Madhulika A. Gupta*, Fiona C. Simpson, Danika C.A. Lyons

Department of Psychiatry, Schulich School of Medicine and Dentistry, University of Western Ontario, London, Ontario, Canada

ARTICLE INFO

Article history:
Received 2 February 2015
Received in revised form
26 June 2015
Accepted 13 July 2015
Available online 3 August 2015

Keywords:
Obstructive sleep apnea
Depression
Anxiety
Psychiatry
CPAP
Positive airway pressure
Excessive daytime sleepiness

SUMMARY

Patients with obstructive sleep apnea (OSA) frequently present with symptoms of depression and anxiety. The objective of this study is to determine if treatment with positive airway pressure (PAP) improves symptoms of depression and anxiety. A systematic review was conducted to identify clinical trials of PAP that contained a validated measure of depression severity. Meta-analysis was conducted for depression, anxiety, excessive daytime sleepiness (EDS), quality of life (QoL) and respiratory variables. The systematic review included 33 reports. Pre-post-test analysis of PAP showed a moderate effect size (Hedge's g, 95% CI) for depression 0.524 [0.401–0.647], but a low effect size compared to oral placebo (0.355 [0.187–0.524]) and no effect when compared to dental appliances (0.107 [−0.72–0.287]) and sham PAP (−0.049 [−0.292–0.194]). Anxiety, EDS, and QoL showed similar improvement in pre-post-test analysis, but a lack of superiority to dental appliances and sham PAP. PAP was superior to all comparators for respiratory variables. PAP has a moderate clinical effect on symptoms of depression and anxiety in OSA, but it is not superior to dental appliances or sham PAP. The improvement in subjective symptoms, such as depression and anxiety, may be mediated by patient expectations and contact with healthcare providers.

© 2015 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Introduction

Obstructive sleep apnea (OSA) is characterized by repetitive sleep-related episodes of complete (apnea) or partial (hypopnea) obstruction of the upper airway or respiratory effort related arousals (RERAs) [1,2]. Globally, OSA associated with daytime sleepiness is reported to occur in 2%–5% of adult women and 3%–7% of adult men [1,2]. OSA is associated with multiple other medical comorbidities including; hypertension, insulin-resistance, type II diabetes, and obesity [2]. Positive airway pressure (PAP) delivered in continuous (CPAP), bilevel (BiPAP) or autotitrating (APAP) modes is a standard therapy for moderate (respiratory disturbance index (RDI) ≥15 and ≤30) to severe (RDI>30/hour) OSA and an optional therapy for mild (RDI≥5 and <15) OSA [3,4]. PAP is believed to be effective in reducing the apnea-hypopnea index (AHI) or RDI by providing pneumatic splinting of the upper airway during sleep [3,5]. There is a large body of evidence showing that PAP improves objective symptoms of OSA including the AHI and blood pressure [6].

The presence of psychological symptoms in patients with OSA has been recognized since the 1970s [7] and significant evidence that there are elevated levels of OSA in patients with clinical depression has emerged since the early 2000s [8]. The recognition of the importance of depression in OSA is exemplified by the fact that in the most recent edition of the International classification of sleep disorders [ICSD3], diagnosis of a mood disorder (Criterion A.4) is among one of the four major groups of symptoms (eg., excessive daytime sleepiness, habitual snoring) or disorders (eg., hypertension, atrial fibrillation) that must be present along with ≥5 obstructive respiratory events per hour, before an OSA diagnosis can be made.

The relationship between objectively-rated features of OSA and symptoms like depression, anxiety, and excessive daytime sleepiness (EDS) is poorly understood [8–10]. This constellation of symptoms is often referred to as subjective symptoms of OSA, as there is typically no consistent correlation between the severity of these symptoms and the severity of OSA measured by objectively-rated sleep variables [11–17]. Current models for the relationship between OSA and psychiatric conditions suggest that underlying biological, metabolic and neurological dysregulation contributes to a feed-forward mechanism that manifests as psychiatric disorders, sleep fragmentation,

* Corresponding author. 585 Springbank Drive, Suite 101 London, Ontario, N6J 1H3, Canada.
E-mail address: magupta@uwo.ca (M.A. Gupta).

http://dx.doi.org/10.1016/j.smrv.2015.07.002
1087-0792/© 2015 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Abbreviations

AHI	apnea hypopnea index
AI	arousal index
APAP	autotitrating positive airway pressure
BAI	Beck anxiety inventory
BDI or BDI-II	Beck depression inventory
BiPAP	bilevel positive airway pressure
BSI-A	brief symptom inventory – anxiety subscale
BSI-D	brief symptom inventory – depression subscale
CES-D	center for epidemiological studies depression scale
CONSORT	consolidated standards of reporting trials
CPAP	continuous positive airway pressure
DA	dental appliance
EDS	excessive daytime sleepiness
ESS	Epworth sleepiness scale
FOSQ	functional outcomes of sleep questionnaire
GHQ-28	general health questionnaire
GRADE	grades of recommendation, assessment, development and evaluation
HADS	hospital anxiety and depression scale
HAM-D	Hamilton rating scale for depression
ICSD3	International classification of sleep disorders 3
ITT	intention-to-treat
MADRS	Montgomery–Åsberg depression rating scale
Mean SaO_2	mean oxygen saturation
Min SaO_2	minimum oxygen saturation
MMPI	Minnesota multiphasic personality inventory
MSLT	multiple sleep latency test
N/A	not applicable
NHP-2	Nottingham health profile-2
OP	oral placebo
OSA	obstructive sleep apnea
PAP	positive airway pressure
POMS	profile of mood states
PRISMA	preferred reporting items for systematic reviews and meta-analyses
QoL	quality of life
RCT-P	parallel-group randomized controlled trial
RCT-X	crossover randomized controlled trial
RDI	respiratory disturbance index
REM%	rapid eye movement sleep percentage
RoB	risk of bias
SAQLI	sleep apnea quality of life index
SAT	single assignment trial
SDS	Zung self rated depression scale
SF-36	short form health survey
SSS	Stanford sleepiness scale
STAI	state trait anxiety inventory
TAS	tension anxiety scale
TEAE	treatment emergent adverse event
WHO-5	World Health Organization-five well-being index
WHOQoL-B	World Health Organization quality of life-brief questionnaire (all subscales)

cardiovascular and metabolic disease [8–10]. Given that the subjective symptoms of OSA are the result of complex biological dysregulation, it remains unclear whether these symptoms have a direct response to PAP therapy. Saunamaki et al. found that symptoms of depression and anxiety were commonly reported in treatment naïve OSA patients with multiple studies reporting improvements in these symptoms after ≥3 mo of PAP [11]. Sanchez et al. also reported on improvements in symptoms of EDS and mood in response to CPAP therapy [18]. Both of these reviews reached their conclusions by focusing on the number of studies reporting the positive effects of PAP; this type of analysis is less accurate as it presents the overall research in the field by study outcome, whereas meta-analysis pools participant data across studies to create a more accurate analysis of the effects of PAP based on the total population studied [11,18]. A recent meta-analysis concluded that treatment of OSA with PAP or mandibular advancement devices is superior to placebo for the treatment of depressive symptoms, but did not evaluate PAP vs. active therapies [19]. The goal of this systematic review and meta-analysis is to determine if PAP therapy improves depression and other subjective symptoms in OSA in comparison to both placebo and active treatments, as measured by validated rating scales.

Objective

The primary objective is to determine if treatment with PAP improves the psychological symptoms of depression in patients with OSA. The secondary objective was to evaluate if PAP improved additional subjective sleep outcomes including: anxiety and EDS.

Methods

The methodology for this paper follows PROSPERO protocol CRD42014007419 registered February 3, 2014 in accordance with the preferred reporting items for systematic reviews and meta-analyses (PRISMA) statement [20,21].

Eligibility criteria

Types of studies

This review considered single arm trials (SAT), case-control trials, and randomized controlled trials (RCTs) of both parallel-group and crossover design of PAP interventions for OSA.

Types of participants

Adult men and women with OSA diagnosed by polysomnography (PSG) were included. OSA was defined as an AHI ≥5 (or RDI equivalent). Studies were only included when participants were treatment naïve at study initiation. Studies where the primary study sample involved individuals with severe or acute co-morbid medical illness such as acute myocardial infarction, stroke, dementia or heart failure were also excluded.

Types of interventions

Studies included PAP devices titrated to an effective pressure to overcome the respiratory disturbance. The analyses include studies with comparators consisting of healthy controls, sham or placebo PAP groups or active comparator controlled studies. This analysis also considered participant controlled pre- and post-intervention data.

Types of outcome measures

The primary outcome measure for the meta-analysis was depression. On this basis, all studies were required to include at least one validated measure of depression such as the Beck depression inventory (BDI), Montgomery-Asberg depression rating scale (MADRS), hospital anxiety and depression scale (HADS), or Hamilton depression scale (HAM-D).

Secondary outcome measures included anxiety, EDS, and quality of life (QoL). This meta-analysis also evaluated the polysomnographic (PSG) indices: AHI/RDI, mean oxygen saturation (Mean SaO_2), minimum oxygen saturation (Min SaO_2), and arousal index (AI). These outcomes were included as they have known physiological efficacy in OSA. As known objective outcome measures, they can provide context for the subjective measures evaluated.

Additional outcome measures present in the studies were included if they were identified during data extraction.

Information sources

Electronic searches

The Cochrane central register of controlled trials (CENTRAL) in The Cochrane Library;

MEDLINE via OVID (from 1946);

EMBASE via OVID (from 1947);

PsycInfo via OVID (from 1806) and

Trials registers

The *meta*Register of controlled trials (www.controlled-trials.com);

The US National Institutes of Health ongoing trials register (www.clinicaltrials.gov);

The Australian New Zealand clinical trials registry (www.anzctr.org.au);

The World Health Organization International clinical trials registry platform (www.who.int/trialsearch);

The EU clinical trials register (www.clinicaltrialsregister.eu);

The Latin American and Caribbean health science information database (http://lilacs.bvsalud.org/en/).

Search

The search strategies for all databases are included in Supplement 1. The reference lists of all primary studies and review articles were checked for additional references.

Study selection

Studies were screened for relevance to the topic by title, followed by abstract and/or full text article. Articles were screened for inclusion by two reviewers (FS and DL). Studies that were not in English or conducted using animals were excluded.

Data collection process

Data extraction was performed by two independent reviewers (FS and DL) using data collection forms. The data forms were piloted before use and included: Study identifier, study design, number of participants per group, baseline characteristics, intervention, treatment regimen and duration, treatment success and failure, efficacy, safety, morbidity and mortality, tolerability, compliance, number of patients lost to follow-up, and the duration of follow-up. Disagreements were resolved with discussion, or if agreement could not be reached, a third author (MAG) assisted in the resolution of any discrepancies. The data were entered by a single author (FS) and verified by both reviewers (FS and DL). Data were entered as intention-to-treat (ITT) data, and per protocol results was converted into ITT data wherever possible to reduce bias.

Risk of bias in individual studies

The risk of bias (RoB) in individual studies was evaluated using the Cochrane risk of bias tool in RevMan version 5.2 (Copenhagen, Denmark) [22,23].

Measures of treatment effect

Treatment effect was assessed by meta-analysis if there were $\geq$2 studies that assessed the outcome measure. Effect size was presented using Hedges' g for continuous measures and odds ratios (OR) for binary data [22,24]. If there was insufficient evidence for meta-analysis, the study outcomes were described verbally, and not included in the meta-analysis.

Unit of analysis issues

The unit of analysis reported was the participant. In cross-over study designs, participant controlled data were used wherever possible or the first intervention period was used where participant-controlled data were unavailable. If studies reported more than one outcome measure included in this review, all of the outcomes were included in the analyses.

Dealing with missing data

Principal investigators and/or study sponsors were contacted to obtain missing data and verify any unclear study parameters. Where this was not possible, and the missing data was felt to introduce a serious bias, sensitivity analysis was conducted to explore the impact of the missing data.

Assessment of heterogeneity

The I^2 statistic was used to measure the proportion of variance attributable to heterogeneity among the trials in each analysis [22,24]. Where substantial heterogeneity was identified, it was reported and possible causes were explored by pre-specified subgroup analysis. Heterogeneity was stratified into I^2 values for low (25%), medium (50%) and high (>75%) heterogeneity [24].

Assessment of reporting biases

Analyses containing $\geq$10 studies were assessed for publication bias using funnel plots and Duval and Tweedie's *trim and fill* method [24].

Data synthesis

All meta-analyses were conducted with a pre-specified random-effects model to account for heterogeneous treatment effects and varying true effect sizes across studies. Meta-analyses were required to have an I^2 value $\leq$75% for inclusion. Where more than one scale was used to assess an outcome measure (i.e., depression by BDI and HADS), the mean of the outcomes was used. Pre-post test analysis used the PAP arm of SATs, parallel RCTs, and cross-over RCTs where pre- and post-test data were presented. The analyses for RCTs used both parallel and crossover designs, where the crossover trials presented participant controlled data. The meta-analysis calculations were performed using Comprehensive Meta-Analysis Version 2.2.064 (Englewood, NJ) [25].

Risk of bias across studies

The RoB across studies was evaluated for each meta-analysis using grades of recommendation, assessment, development and evaluation (GRADE) and GRADEpro software [22,26].

Subgroup analysis and meta-regression

Prospective subgroup and meta-regression analysis were intended to be conducted by gender, OSA severity (AHI/RDI), body

mass index (BMI), arousal index (AI), and baseline depression severity based on the heterogeneity of the results. However, as all of the pre-specified analyses had heterogeneity below 50%, these analyses were not included in the final manuscript.

Results

Description of studies

Results of search

The results of the search strategy can be found in Fig. 1. The search was completed on June 20, 2014. Thirty-three reports of 31 studies met the inclusion criteria for this review (Supplement, Table S2) [27–59].

Included Studies

A list of the included studies can be found in Table S2. The characteristics of the included studies are summarized in Table 1.

Study design

The included studies are comprised of single assignment trials, randomized trials, and current or future study protocols (Table 1). Fig. 1 shows the number of each class of trial identified. Of the randomized studies, four were parallel assignment studies [42,45,46,52,59] and six were crossover trials [28,37–40,47].

Participants

The included studies enrolled a total of 895 participants. Seven studies enrolled fewer than 20 participants [33,34,39,48,50,52,58], and two studies enrolled more than 100 participants [28,44]. Participant mean BMI ranged from 27.8 to 38 kg/m^2 where recorded, indicating that the participants ranged from overweight to obese [34,52,58]. Participant mean baseline AHI ranged from 11.0 to 71.5 per hour where reported, including a full spectrum of mild to severe OSA [39,48].

Interventions

A total of seven interventions were measured in the included studies. All of the studies evaluated CPAP devices. In addition to the required PAP treatment, the completed studies included oral placebo [28,37–40], oral/dental appliances [28,52], exercise [52], and sham PAP [42,45–47,59]. Two trial protocols also included gastric bypass surgery and lifestyle/behavioural modification as arms of parallel-assignment trials [30,31]. The studies ranged in duration from 11 d to two y [57,59].

Outcomes

The included studies reported a diverse array of psychological and physiological outcome measures (Table 2). EDS was reported as an outcome measure for 18 completed trials [28,32,33,35–41,43,44,47,48,52–54,57,58] and five trial protocols

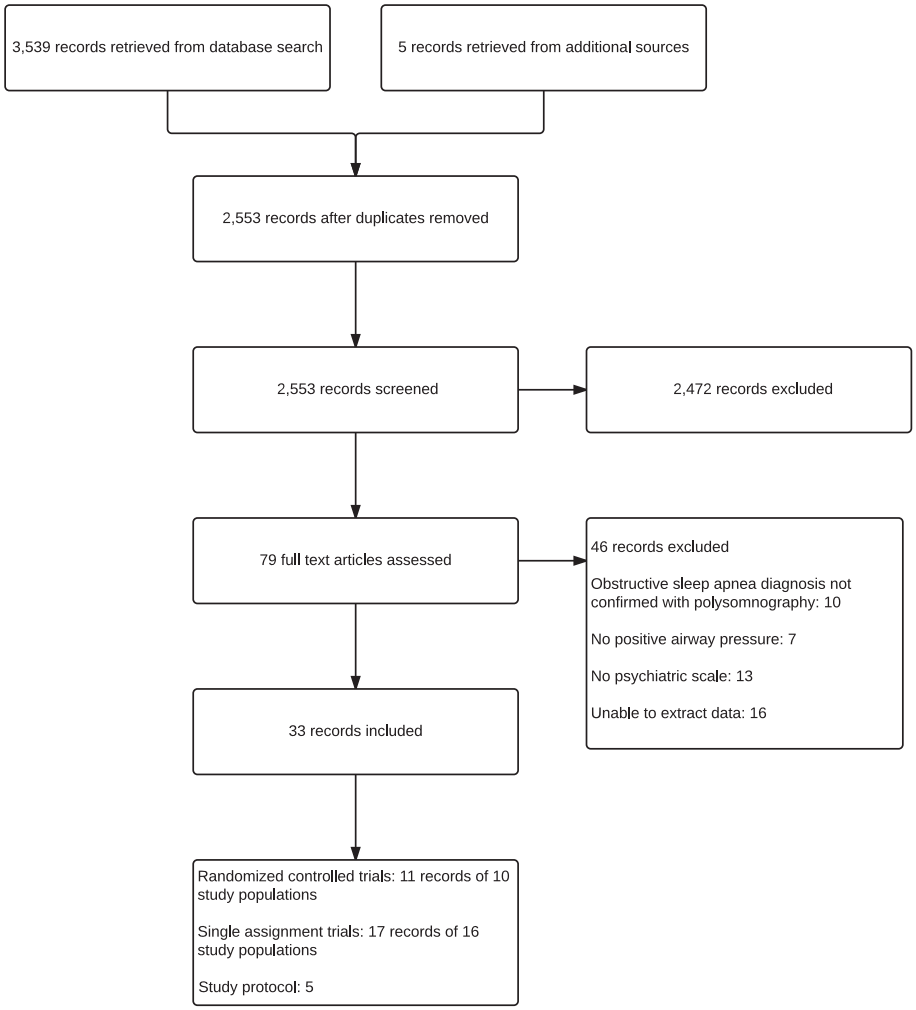


Fig. 1. Flow diagram of systematic review. The systematic review resulted in 33 reports of 31 studies. In cases where there were multiple records of a study population, they were combined under a single study identifier to prevent duplication.

Table 1
Summary of included studies.

Study	Trial type	Intervention	Control	ITT sample	Study duration	Psychiatric scales	EDS scale	QoL scale	Age (mean ± SD)	Male (%)	BMI (kg/m^2, mean ± SD)	AHI/RDI (events/h, mean ± SD)
AFIP 2014 [27]	Protocol	CPAP	Oral appliance	90*	3 y	BDI, BAI	ESS	–	N/A	N/A	N/A	N/A
Barnes 2004 [28]	RCT-X	CPAP	Dental appliance or oral placebo	114	3 mo	BDI, POMS	ESS, SSS	SF-36	47 ± 0.9	80%	31.1 ± 0.5	21.3 ± 1.3
Borak 1996 [29]	SAT	CPAP	–	20	12 mo	BDI, TAS	–	–	46 ± 6	–	–	67 ± 16
BWH 2014 [30]	Protocol	CPAP	Gastric bypass surgery	80*	9 mo	PHQ-9	ESS	GQL	N/A	N/A	N/A	N/A
Campos-Rodriguez 2014 [31]	Protocol	CPAP	Dietary and lifestyle advice	300*	12 wk	POMS, HADS	ESS	SF-12,	N/A	N/A	N/A	N/A
Castronovo 2009 [32,33]	SAT	CPAP	–	14	4 mo	BDI	ESS	SF-36	43.93 ± 7.8	100%	30.29 ± 4.8	50.14 ± 24.8
Derderian 1988 [34]	SAT	CPAP	–	7	2 mo	POMS	–	–	59.3 ± 5.8	100%	27.8 ± 2.4	–
Diamanti 2013 [35]	SAT	CPAP	–	24	8 mo	CES-D	QSQ	WHOQoL	51.88 ± 10.6	83%	34.37 ± 6.5	37.5 ± 6.45
El-Sherbini 2011 [36]	SAT	CPAP	–	37	3 mo	HAM-D	ESS	–	44.9 ± 8.8	65%	–	27.7 ± 7.3
Englemann 1994 [38]	RCT-X	CPAP	Oral placebo	32	4 wk	HADS	ESS	NHP-2	49 ± 8.4	81%	33 ± 1.6	28
Englemann 1997 [39]	RCT-X	CPAP	Oral placebo	16	4 wk	HADS	ESS	GHQ-28, NHP-2	52 ± 7.7	75%	29.8 ± 7	11 ± 3.9
Englemann 1998 [40]	RCT-X	CPAP	Oral placebo	23	4 wk	HADS	ESS	GHQ-28, NHP-2	47 ± 12	91%	30 ± 7	43 ± 37
Englemann 1999 [37]	RCT-X	CPAP	Oral placebo	34	4 wk	HADS	ESS	SF-36, NHP-2	44 ± 8	62%	30 ± 5	24 ± 10
Ferini-Strambi 2003 [41]	SAT	CPAP	–	23	6 mo	BDI	ESS	–	56.52 ± 6.1	91%	33.45 ± 5.6	55 ± 13.4
Haensel 2007 [42]	RCT-P	CPAP	Sham CPAP	25/25	14 d	POMS	–	–	48.2 ± 10.2	80%	33.1 ± 8.2	63.6 ± 29.1
Herold 2011 [43]	SAT	CPAP	–	60	6 mo	BDI, HAM-D, MADRS	ESS, PSQI	WHO-5	54.77 ± 10	38%	–	30 ± 21.9
Kawahara 2005 [44]	SAT	CPAP	–	132	56 d	SDS	ESS	–	48.8 ± 11.9	94%	28.3 ± 4.3	59.4 ± 23.8
Lee 2012 [45,46]	RCT-P	CPAP	Sham CPAP	26/30	21 d	BSI, CES-D, POMS	–	–	46 ± 9.4	84%	29.2 ± 4.2	33.9 ± 20.1
Marshall 2005 [47]	RCT-X	CPAP	Sham CPAP	31	8 wk	HADS	ESS	FOSQ, SF-36	50.5 [25–67]	76%	31.5 ± 6.0	21.6 ± 7.5
O'Donoghue 2012 [48]	SAT	CPAP	–	30	6 mo	BDI	ESS	FOSQ	45.2 ± 9.6	100%	33.3 ± 4.6	71.5 ± 16.2
Paparigopoulos 2009 [49]	SAT	CPAP		56	1 mo	HADS	–	SF-36	47.4 ± 8.1	93%	31 ± 5.7	54.5 ± 27.4
Ramos Platon 1992 [50]	SAT	CPAP	–	23	11 mo	MMPI	–	–	53.3 ± 10.7	97%	–	–
Sanchez 2001 [51]	SAT	CPAP		51	3 mo	BDI, STAI	–	–	48.04 ± 8.6	88%	–	61.32 ± 21.9
Schutz 2013 [52]	RCT-P	CPAP	Dental appliance or exercise	9/9/7	60 d	POMS	ESS	SF-36	41.1 ± 7.5	–	27.8 ± 3.5	26.2 ± 14.7
Schwartz 2005 [54]	SAT	CPAP	–	50	4-6 wk	BDI	ESS	–	48.4 ± 9.3	66%	37 ± 8.6	55.8 ± 34.4†
Schwartz 2007 [53]	SAT	CPAP	–	50	600 d	BDI	ESS	–	53 ± 11.3	78%	35 ± 7.7	57.4 ± 31.1†
Stanford 2014 [55]	Protocol	CPAP	Oral appliance	238*	6-9 mo	POMS	FOSQ, SAQLI	SF-36	N/A	N/A	N/A	N/A
Uppsala 2014 [56]	Protocol	CPAP	Physical activity to promote weight loss	140*	18 mo	MADRS	ESS	SF-36	N/A	N/A	N/A	N/A
Yamamoto 2000 [57]	SAT	CPAP	–	46	2 y	SDS	ESS	–	49.5 ± 10.8	100%	29.2 ± 5.4	–
Ye 2009 [58]	SAT	CPAP	–	176	3 mo	POMS	ESS, FOSQ, MSLT	–	46.7 ± 8.8	86%	38 ± 8.2	63.9 ± 29.4
Yu 1999 [59]	RCT-P	CPAP	Sham CPAP	20/14	11 d	POMS	–	–	48.2 ± 8.9	74%	–	40.55 ± 23.6†

(–) not reported, (*) Planned enrollment, (†) RDI - respiratory disturbance index, N/A – not applicable, AHI – apnea hypopnea index, BAI – Beck anxiety inventory, BDI – Beck depression inventory, CPAP – continuous positive airway pressure, CES-D – Center for epidemiological studies depression scale, ESS – Epworth sleepiness scale, FOSQ – functional outcomes of sleep questionnaire, GHQ-28 – general health questionnaire, HADS – hospital anxiety and depression scale, HAM-D – Hamilton rating scale for depression, ITT – intention-to-treat, MADRS – Montgomery-Åsberg depression rating scale, MMPI – Minnesota multiphasic personality inventory, MSLT – multiple sleep latency test, POMS – profile of mood states, RCT-P – parallel-group randomized controlled trial, RCT-X – crossover randomized controlled trial, SAQLI – sleep apnea quality of life index, SAT – single assignment trial, SDS – Zung self-rated depression scale, SSS – Stanford sleepiness scale.

Table 2
Outcome measures.

Outcome measure	Scales assessed
Depression	BDI or BDI-II – Beck depression inventory, BSI-D – brief symptom inventory – depression subscale, CES-D – Center for epidemiological studies depression scale, HADS-D – hospital anxiety and depression scale – depression subscale, HAM-D – Hamilton rating scale for depression, MADRS – Montgomery-Åsberg depression dating scale, MMPI – Minnesota multiphasic personality inventory – depression subscale, POMS-D – profile of mood states – depression subscale, SDS – Zung self-rated depression scale
Excessive daytime sleepiness (EDS)	ESS – Epworth sleepiness scale, QSQ- Quebec sleepiness questionnaire , SSS – Stanford sleepiness scale
Anxiety	BSI-A – brief symptom inventory – anxiety subscale, HADS-A – hospital anxiety and depression scale – anxiety subscale, MMPI-Pt – Minnesota multiphasic personality inventory- psychasthenia subscale, POMS-T – profile of mood states -tension and anxiety subscale, STAI – state trait anxiety inventory, TAS – tension anxiety scale
Quality of life (QoL)	GHQ-28 – general health questionnaire, NHP-2 – Nottingham health profile-2, SF-36 - short form health survey (all subscales), WHO-5 – WHO-five well-being index , WHOQoL-B – World Health Organization auality of life-brief questionnaire (all subscales)

[27,30,31,55,56]. An anxiety scale or subscale result was reported by 15 completed trials [28,29,34,37–40,42,45–47,49,50,52,58,59] and three trial protocols [27,31,55]. QoL was reported by eight completed trials [28,32,33,35,37,39,40,43,47,49,52] and four trial protocols [30,31,55,56].

Funding

Thirteen of the included studies had no identifiable source of funding [34–36,38,41,43,44,49–51,53,54,57]. Twelve of the studies received funding from government grants or charitable organizations [28,29,31–33,37,40,42,45–47,52,58,59]. Four studies received equipment or drugs from commercial interests [28,39,48,58].

Excluded studies

There were 46 studies excluded from the full text search (Table S3). The primary reason for exclusion was an inability to extract data from studies that otherwise met the inclusion criteria, when contacting the authors failed (Table S4). The secondary reasons were lack of a psychiatric scale and OSA diagnoses reported without PSG results (Table S4).

Risk of bias in included studies

Allocation

The studies identified in this systematic review were primarily single-assignment prospective trials or retrospective reviews. Of the 10 completed randomized controlled studies, two had adequate randomization procedures (Fig. 2) [37,45]. Allocation concealment was only possible in the four sham controlled PAP studies. The method of concealment was not described in three of these studies [42,45,59] and in the fourth study participants were informed that the trial was for two different pressures of humidified PAP [47].

Blinding

The blinding of studies in this systematic review was subject to either unclear or high risk of bias (Fig. 2). The single-assignment trials largely do not discuss blinding and are presumably open label trials. Of the 10 randomized controlled trials, five are crossover studies of mixed treatments which cannot be participant-blinded [28,37–40], one is a crossover of PAP and sham PAP which does not report blinding [47], and two of these are open-label trials without evaluator blinding [38,40]. Of the five parallel group studies, two had adequate blinding of participants and evaluators [45–47], while the other three had unclear blinding procedures [42,52,59].

Incomplete outcome data

Incomplete outcome data were absent in the majority of completed studies (Fig. 2). There was a low risk of bias for incomplete outcome data in 20 studies. Unclear outcome data were present in two studies [29,49]. There was a high risk of bias for incomplete outcome data in three studies [44,54,59].

Effect of interventions

The magnitude of change caused by the interventions is presented as effect size measured as Hedge's g. The interpretation of these effect sizes requires context to determine how the statistical effect size corresponds to the magnitude of clinical change observed. Based on the range of effect sizes presented in this analysis and their correlations to known objective measures of the effects of PAP on sleep symptoms such as AHI [6], effect sizes will be interpreted as <0.5 low, 0.5–1.0 moderate and >1.0 as high.

PAP pre-post intervention

Twenty-one of the included studies assessed the change in depression and other outcome measures pre- and post-PAP treatment (Table 3) [28,29,33–37,40–44,47–52,54,57,59]. Comparison of participants who underwent PAP therapy at baseline and follow-up shows positive effect sizes for all outcome measures post-treatment (Table 3). These analyses are subject to an overall high RoB with a GRADE evaluation of 'very low' for all outcome measures other than AHI/RDI which was 'low'. The effect size for the primary outcome measure of depressive symptoms was 0.524 [0.401–0.647], which corresponds to a moderate clinical effect (Fig. 3A). A high effect size was observed for symptoms of EDS (1.015 [0.697–1.333]) (Fig. 3C). The secondary outcome measures of anxiety (0.413 [0.263–0.563], Fig. 3B) and QoL (0.435 [0.236–0.634], Fig. 3D) had low effect sizes. The FOSQ has a moderate outcome (0.509 [0.219–0.799], Figure S1D). The objective sleep measure AHI/RDI had a large effect size (1.479 [1.277–1.680]) (Figure S1A). The AI (0.624 [−0.137–1.384], Figure S1C) and mean SaO_2 (0.972 [−0.159–2.104], Figure S1B) also showed moderate effect sizes, but the 95% CI indicates that these are not clinically significant.

PAP versus oral placebo

Comparisons of PAP with oral placebo were conducted in five of the included studies (Table 4, Fig. 4) [28,37–40]. All of the analyses indicate that PAP is more effective than oral placebo (Table 4, Fig. 4 and S2). These analyses are subject to an overall high RoB with a GRADE evaluation of 'low' for all outcome measures other than EDS which was 'moderate'. The primary outcome measure for depression showed a low effect size of 0.355 [0.187–0.524] (Fig. 4A). The change in EDS was moderate with an effect size of 0.608 [0.090–1.126] (Fig. 4C). Anxiety (0.225 [0.093–0.357], Fig. 4B) and QoL (0.491 [0.126–0.855], Fig. 4D) had low effect sizes favoring PAP. The meta-analysis for QoL had moderate heterogeneity of 33.96% which is larger than the heterogeneity observed for the other analyses in this comparison. The moderate QoL heterogeneity

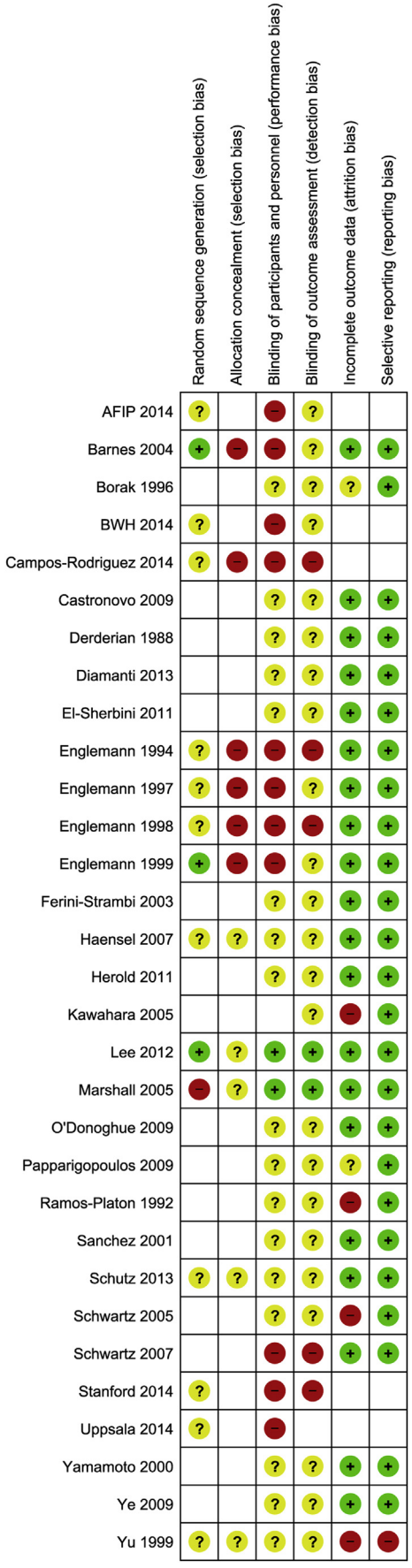


Fig. 2. Individual RoB for the included studies. (+) low risk of bias, (?) unclear risk of bias, (−) high risk of bias.

between the included studies may be due to the number of varying QoL scales that were utilized.

PAP versus dental appliances

Two studies compared PAP and dental appliances (Table 5, Fig. 5 and S2) [28,52]. The RoB for both studies was 'low'. The comparison between PAP and dental appliances did not show a significant difference between the two treatments for depression (0.107 [−0.72–0.287], Fig. 5A), EDS (0.058 [−0.121–0.237], Fig. 5C) or QoL (0.021 [−0.157–0.200], Fig. 5D). There was a low effect size favoring PAP for anxiety 0.101 [0.078–0.280] (Fig. 5B). The effect size for AHI/RDI favored PAP over dental appliances with a moderate effect size of 0.902 [0.692–1.113] (Figure S2A). There was also a low effect size of 0.473 [0.285–0.661] favoring PAP for improvements in AI (Figure S2B).

PAP versus sham PAP

Four studies compared PAP to sham PAP (Table 6, Fig. 6 and S4) [42,45–47,59]. The analyses for depression, anxiety and mean SaO_2 received 'low' GRADE assessment scores, while the AHI/RDI analysis received a 'high' GRADE assessment. The effect sizes for depression (−0.049 [−0.292–0.194], Fig. 6A) and anxiety (−0.073 [−0.315–0.169], Fig. 6B) show that PAP is not more effective than sham PAP at reducing psychological symptoms. There was a high effect size for AHI/RDI in favor of PAP 1.881 [1.370–2.391] (Figure S3A). The effect size for mean SaO_2 was 0.575 [−0.420–1.570] (Figure S3B), but the large 95% CI indicates that this result has too much variance to be considered clinically relevant. EDS and QoL were not assessed in the PAP vs. sham PAP studies.

PAP vs. exercise

One study evaluated PAP vs. exercise for the treatment of OSA [52]. Meta-analysis was not possible; however, this study did measure the difference in depression, EDS, anxiety, QoL and AHI. All of the effect sizes favored PAP over exercise. Depression and anxiety were measured using the profile of mood states (POMS) scale. The effect size for POMS-depression was low at 0.094 ± 0.477 (Hedge's g ± SE) and the effect size for POMS-tension and anxiety was also low at 0.349 ± 0.480. EDS was measured by the Epworth sleepiness scale (ESS) with a moderate effect size of 0.709 ± 0.493. The effect sizes for respiratory variables were both large with 2.007 ± 0.594 for AHI and 1.033 ± 0.510 for AI. The effect size for the SF-36 mean score was also large at 1.225 ± 0.523.

Discussion

Summary of main results

Our systematic review identified 31 studies pertaining to the effects of PAP on depressive symptoms in individuals with OSA. Of the included studies, 26 were completed trials and five were trial protocols that are currently active or enrolling participants. The majority of studies were single-assignment prospective trials or retrospective reviews. The results of this meta-analysis show that PAP had a moderate impact on symptoms of depression and anxiety when the psychological measures were carried out pre-to post-PAP. PAP had low impact on symptoms of depression and anxiety when compared with oral placebo, but was not more effective at reducing depression and anxiety than sham PAP or dental appliances. In contrast, PAP was consistently more effective than placebo, sham PAP or dental appliances in decreasing the AHI.

This review confirms that the relationship between depression, anxiety, EDS and OSA is complex. The contrast between subjective and objective symptoms is underscored by the fact that PAP is consistently more effective than placebo and active treatment

Table 3
Pre-versus Post-PAP.

Outcome measure	No. of studies	No. of participants	Statistical method	Effect size	Heterogeneity (I^2)	GRADE assessment
Depression	21	803	Hedge's g (Random, 95% CI)	0.524 [0.401–0.647]	4.24%	Very low
Anxiety	12	406	Hedge's g (Random, 95% CI)	0.413 [0.263–0.563]	3.62%	Very low
EDS	16	832	Hedge's g (Random, 95% CI)	1.015 [0.697–1.333]	0%	Very low
QoL	9	365	Hedge's g (Random, 95% CI)	0.435 [0.236–0.634]	4.32%	Very low
AHI/RDI	6	278	Hedge's g (Random, 95% CI)	1.479 [1.277–1.680]	5.54%	Low
Mean SaO_2	3	52	Hedge's g (Random, 95% CI)	0.972 [−0.159–2.104]	0%	Very low
AI	2	123	Hedge's g (Random, 95% CI)	0.624 [−0.137–1.384]	0%	Very low
FOSQ	4	334	Hedge's g (Random, 95% CI)	0.509 [0.219–0.799]	49.89%	Very low

AHI – apnea hypopnea index, AI - arousal index, EDS - excessive daytime sleepiness, FOSQ - functional outcomes of sleep questionnaire, GRADE - grades of recommendation, assessment, development and evaluation, mean SaO_2 – mean oxygen saturation, QoL – quality of life, RDI - respiratory disturbance index.

comparisons at improving all respiratory variables; including AHI/RDI, AI, and mean SaO_2. It has previously been hypothesized that OSA, psychological symptoms, and metabolic disease are the diverse consequences of underlying biological, metabolic and neurologic dysregulation [8–10]. Psychiatric, sleep and metabolic disorders can each be the primary manifestation and entry point into a feed-forward mechanism where underlying inflammation, oxidative and nitrosative stress, and neurotransmitter imbalances increase the likelihood of the individual expressing multiple disease states related to the underlying biological dysregulation [8]. In some individuals, the development of OSA and resultant sleep fragmentation may unmask an underlying predisposition for depressive disorders. The application of PAP is unlikely to resolve these problems directly, as PAP can only directly treat the respiratory symptoms that are an outcome of a complex process. Improvement in AHI that is dependant solely on PAP, without other autonomic and metabolic changes, may not be sufficient to be associated with a clinically significant improvement in depression, anxiety and EDS scores. The improvements in these symptoms may derive from decreased arousals and hypoxemia or the long-term positive impact of PAP on metabolic function [60,61]. Alternatively, the improvement in depressive symptoms in pre-post-test analysis and in comparison to oral placebo may be the result of patient expectations, mediated primarily by interaction with the healthcare system and reassurances of improvement in the participant's condition [62]. Many of the included studies occur over prolonged durations >90 d, where this interaction may have a significant impact.

Overall completeness and applicability of evidence

The evidence for the efficacy of OSA in alleviating symptoms of depression is largely incomplete. The duration of the studies ranged from 11 d to 2 y (Table 1) and this is likely an important confounding factor in the improvement of a complex syndrome like depression after treatment of the respiratory component of OSA with PAP therapy. The lack of significant heterogeneity in the analyses suggests that time was not a significant factor in these analyses. One possible reason for the lack of an impact is that PAP is predominantly a symptomatic treatment that is only effective when being used, so the length of treatment is not a significant factor, despite known improvements in cardiovascular and endocrine function with PAP [6,60,61].

A second concern is the lack of RCTs assessing the effects of PAP in comparison to dental appliances, sham PAP and exercise. While there is an abundance of low quality evidence for the effect of PAP pre- and post-treatment, there is insufficient evidence for the assessment of the outcomes of PAP vs. placebo and active comparators. The highest volume of data exists for the comparison of PAP vs. oral placebo, but this is likely the least convincing placebo arm for participants, despite instructions to the contrary. The evidence for comparing PAP to dental appliances for psychological symptoms is sparse, as is the evidence for PAP vs. sham PAP which is the gold standard control.

An additional problem with the completeness of this evidence is the lack of dichotomous outcomes for clinical depression and anxiety. This meta-analysis exclusively uses continuous outcome measures of varying levels of psychological symptom severity. The combination of multiple scales measuring the same psychometric properties, but on different numerical scales, prevents us from conducting meta-regression to determine if baseline severity has an effect on the magnitude of change in depression. It would be valuable to have data organized to give the proportion of participants meeting the severity criteria for clinical depression before and after PAP to see if these proportions change. It would also be valuable to be able to perform subgroup analysis on continuous outcomes within participants grouped by severity.

Quality of the evidence

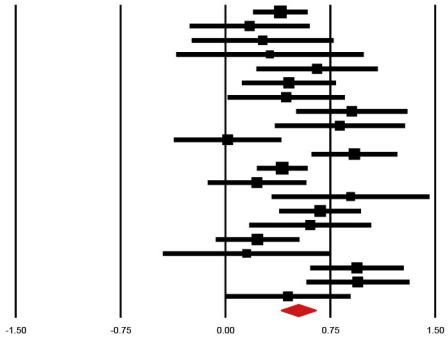
The quality of the evidence included in this review and meta-analysis is low overall, reflecting the fact that the bulk of the data was published before the widespread adoption of the consolidated standards of reporting trials (CONSORT) statement, which is based on the two-group parallel design [63]. The GRADE assessments for the meta-analyses performed in this review were predominantly 'very low'. The decision to evaluate single-arm trials of PAP was based largely on the lack of availability of randomized controlled trials; however, the quality of the evidence based upon single-arm trials should be taken into consideration when interpreting the clinical efficacy indicated by the results. It is particularly telling that the effect size for the primary outcome measure of depression diminishes as the methodological quality of the RCTs comparator arm improves. This effect can be underscored by the fact that the effect size for depression is null −0.049 [−0.292–0.194] when PAP is compared to sham PAP, but the effect size for change in AHI/RDI for this comparison is the largest effect size 1.881 [1.370–2.391] reported in the analysis.

Limitations

This meta-analysis includes single-assignment and randomized clinical trials with both parallel group and cross-over designs, but there are more SATs than RCTs present in the analysis. A similar meta-analysis published assessing mandibular advancement devices or PAP in comparison to placebo focused solely on RCTs [19]. There are advantages to pre-post PAP analysis, but the low RoB associated with SATs cannot be overlooked. The other limitation of this meta-analysis is the use of multiple scales to measure the same psychometric variable [19]. The use of Hedge's g should help to account for variability in the scale ranges, but this is predicated on each scale having equivalent ability to measure the same underlying construct, which may not be true in all cases. The ratings of depression were primarily participant-rated and

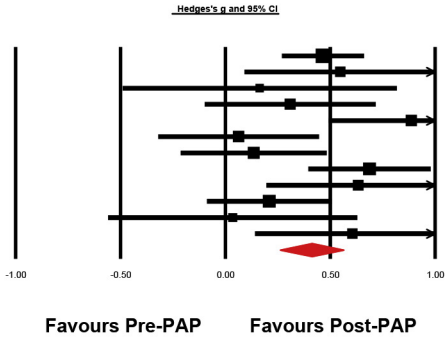
A)

Study name	Subgroup	Comparison	Outcome	Hedges's g	Standard error	Variance	Lower limit	Z-Value	Upper limit	Time point	p-Value	Total	Relative weight
Barnes 2004	Total	Pre-Post PAP	Combined	0.393	0.097	0.009	0.204	4.065	0.583	90	0.000	114	7.29
Borak 1996	Total	Pre-Post PAP	BDI	0.174	0.216	0.047	-0.250	0.805	0.598	365	0.421	20	4.30
Castronovo 2009	Total	Pre-Post PAP	BDI	0.266	0.257	0.066	-0.236	1.039	0.769	90	0.299	14	3.56
Derderian 1988	Total	Pre-Post PAP	POMS-D	0.319	0.340	0.115	-0.347	0.939	0.984	60	0.348	7	2.46
Diamanti 2013	Total	Pre-Post PAP	CES-D	0.656	0.219	0.048	0.227	2.998	1.085	240	0.003	24	4.25
El-Sherbini 2011	Total	Pre-Post PAP	HAM-D	0.454	0.169	0.029	0.122	2.681	0.786	60	0.007	37	5.37
Engleman 1998	Total	Pre-Post PAP	HADS-D	0.435	0.211	0.045	0.021	2.059	0.849	28	0.040	23	4.41
Englemann 1999	Total	Pre-Post PAP	HADS-D	0.904	0.200	0.040	0.511	4.514	1.296	28	0.000	34	4.65
Ferini-Strambi 2003	Total	Pre-Post PAP	BDI	0.819	0.235	0.055	0.359	3.489	1.279	120	0.000	23	3.95
Haensel 2007	Total	Pre-Post PAP	POMS-D	0.017	0.194	0.038	-0.363	0.087	0.397	14	0.930	25	4.79
Herold 2011	Total	Pre-Post PAP	Combined	0.923	0.154	0.024	0.621	5.999	1.224	180	0.000	60	5.77
Kawahara 2005	Total	Pre-Post PAP	SDS	0.407	0.090	0.008	0.230	4.514	0.583	56	0.000	132	7.47
Marshall 2005	Total	Pre-Post PAP	HADS-D	0.226	0.177	0.031	-0.122	1.275	0.574	21	0.202	31	5.17
O'Donoghue 2012	Total	Pre-Post PAP	BDI	0.896	0.285	0.081	0.336	3.139	1.455	180	0.002	16	3.12
Paparrigopoulos 2009	Total	Pre-Post PAP	HADS-D	0.678	0.147	0.021	0.391	4.626	0.965	30	0.000	56	5.96
Ramos Platon 1992	Total	Pre-Post PAP	MMPI-D	0.607	0.220	0.049	0.175	2.755	1.039	210	0.006	23	4.23
Sanchez 2001	Total	Pre-Post PAP	BDI	0.230	0.150	0.023	-0.064	1.531	0.524	30	0.126	44	5.86
Schutz 2013	Total	Pre-Post PAP	POMS-D	0.153	0.303	0.092	-0.442	0.503	0.747	60	0.615	9	2.88
Schwartz 2005	Total	Pre-Post PAP	BDI	0.942	0.168	0.028	0.612	5.603	1.271	30-42	0.000	50	5.40
Yamamoto 2000	Total	Pre-Post PAP	SDS	0.948	0.186	0.034	0.584	5.108	1.312	730	0.000	41	4.98
Yu 1999	Total	Pre-Post PAP	POMS-D	0.448	0.226	0.051	0.005	1.980	0.891	11	0.048	20	4.11
				0.524	0.063	0.004	0.401	8.349	0.647		0.000	803	



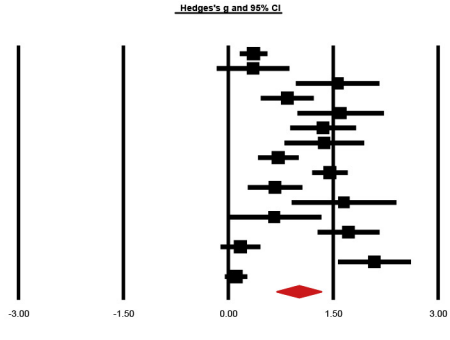
B)

Study name	Subgroup	Comparison	Outcome	Hedges's g	Standard error	Variance	Lower limit	Z-Value	Upper limit	Time point	p-Value	Total	Relative weight
Barnes 2004	Total	Pre-Post PAP	POMS-T	0.465	0.098	0.010	0.273	4.744	0.657	90	0.000	114	14.26
Borak 1996	Total	Pre-Post PAP	TAS	0.548	0.231	0.054	0.094	2.367	1.002	365	0.018	20	6.89
Derderian 1988	Total	Pre-Post PAP	POMS-T	0.164	0.332	0.110	-0.486	0.495	0.814	60	0.621	7	4.14
Engleman 1998	Total	Pre-Post PAP	HADS-A	0.309	0.206	0.043	-0.095	1.498	0.714	28	0.134	23	7.91
Englemann 1999	Total	Pre-Post PAP	HADS-A	0.887	0.199	0.040	0.496	4.453	1.277	28	0.000	34	8.24
Haensel 2007	Total	Pre-Post PAP	POMS-T	0.063	0.194	0.038	-0.317	0.324	0.443	14	0.746	25	8.48
Marshall 2005	Total	Pre-Post PAP	HADS-A	0.135	0.176	0.031	-0.210	0.768	0.480	21	0.442	31	9.38
Paparrigopoulos 2009	Total	Pre-Post PAP	HADS-A	0.687	0.147	0.022	0.399	4.676	0.975	30	0.000	56	11.04
Ramos Platon 1992	Total	Pre-Post PAP	MMPI-PI	0.634	0.222	0.049	0.199	2.855	1.069	210	0.004	23	7.26
Sanchez 2001	Total	Pre-Post PAP	Combined	0.209	0.150	0.023	-0.085	1.395	0.503	30	0.163	44	10.85
Schutz 2013	Total	Pre-Post PAP	POMS-T	0.035	0.301	0.091	-0.555	0.117	0.625	60	0.907	9	4.79
Yu 1999	Total	Pre-Post PAP	POMS-T	0.606	0.235	0.055	0.145	2.576	1.066	11	0.010	20	6.76
				0.413	0.077	0.006	0.263	5.402	0.563		0.000	406	



C)

Study name	Subgroup	Comparison	Outcome	Hedges's g	Standard error	Variance	Lower limit	Z-Value	Upper limit	Time point	p-Value	Total	Relative weight
Barnes 2004	Total	Pre-Post PAP	Combined	0.361	0.096	0.009	0.173	3.761	0.550	90	0.000	114	6.98
Castronovo 2009	Total	Pre-Post PAP	ESS	0.355	0.260	0.068	-0.155	1.363	0.865	90	0.173	14	6.04
Diamanti 2013	Total	Pre-Post PAP	QSQ-DS	1.564	0.300	0.090	0.977	5.216	2.152	240	0.000	24	5.75
El-Sherbini 2011	Total	Pre-Post PAP	ESS	0.844	0.188	0.036	0.474	4.476	1.213	60	0.000	37	6.53
Engleman 1998	Total	Pre-Post PAP	ESS	1.607	0.311	0.097	0.997	5.168	2.216	28	0.000	23	5.67
Englemann 1999	Total	Pre-Post PAP	ESS	1.355	0.235	0.055	0.895	5.774	1.815	28	0.000	34	6.22
Ferini-Strambi 2003	Total	Pre-Post PAP	ESS	1.372	0.285	0.081	0.813	4.807	1.931	120	0.000	23	5.86
Herold 2011	Total	Pre-Post PAP	ESS	0.714	0.143	0.020	0.434	4.990	0.995	180	0.000	60	6.78
Kawahara 2005	Total	Pre-Post PAP	ESS	1.453	0.124	0.015	1.209	11.677	1.697	56	0.000	132	6.87
Marshall 2005	Total	Pre-Post PAP	ESS	0.668	0.195	0.038	0.287	3.434	1.049	21	0.001	31	6.49
O'Donoghue 2012	Total	Pre-Post PAP	ESS	1.654	0.377	0.142	0.916	4.392	2.392	180	0.000	16	5.17
Schutz 2013	Total	Pre-Post PAP	ESS	0.658	0.339	0.115	-0.006	1.944	1.322	60	0.052	9	5.45
Schwartz 2005	Total	Pre-Post PAP	ESS	1.720	0.221	0.049	1.286	7.772	2.154	30-42	0.000	50	6.32
Schwartz 2007	Total	Pre-Post PAP	ESS	0.174	0.140	0.020	-0.101	1.242	0.449	30-42	0.214	50	6.79
Yamamoto 2000	Total	Pre-Post PAP	ESS	2.089	0.262	0.068	1.576	7.985	2.602	730	0.000	46	6.03
Ye 2009	Combined	Pre-Post PAP	ESS	0.111	0.076	0.006	-0.039	1.446	0.261	90	0.148	169	7.05
				1.015	0.162	0.026	0.697	6.256	1.333		0.000	832	



D)

Study name	Subgroup	Comparison	Outcome	Hedges's g	Standard error	Variance	Lower limit	Z-Value	Upper limit	Time point	p-Value	Total	Relative weight
Barnes 2004	Total	Pre-Post PAP	SF-36	0.350	0.096	0.009	0.162	3.647	0.538	90	0.000	114	15.59
Castronovo 2009	Total	Pre-Post PAP	SF-36	0.168	0.254	0.064	-0.329	0.664	0.665	90	0.507	14	8.50
Diamanti 2013	Total	Pre-Post PAP	Combined	0.489	0.211	0.045	0.076	2.319	0.903	240	0.020	24	10.16
Engleman 1998	Total	Pre-Post PAP	Combined	0.329	0.207	0.043	-0.078	1.585	0.735	28	0.113	23	10.31
Englemann 1999	Total	Pre-Post PAP	Combined	0.647	0.188	0.035	0.279	3.449	1.015	28	0.001	34	11.19
Herold 2011	Total	Pre-Post PAP	WHO-5	0.899	0.152	0.023	0.602	5.931	1.196	180	0.000	60	12.90
Marshall 2005	Total	Pre-Post PAP	Combined	0.064	0.177	0.031	-0.282	0.364	0.410	21	0.716	31	11.70
Paparrigopoulos 2009	Total	Pre-Post PAP	Combined	0.712	0.149	0.022	0.420	4.772	1.005	30	0.000	56	13.02
Schutz 2013	Total	Pre-Post PAP	Combined	-0.095	0.314	0.099	-0.711	-0.303	0.520	60	0.762	9	6.63
				0.435	0.102	0.010	0.236	4.282	0.634		0.000	365	

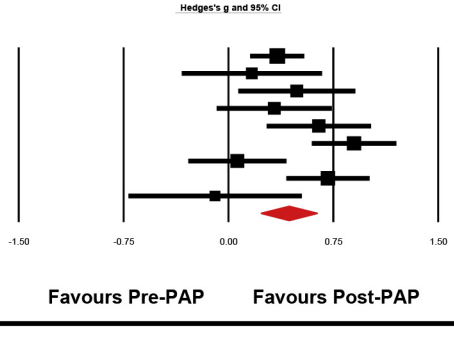


Fig. 3. Pre-post PAP Forest plots. A) Depression, B) Anxiety, C) EDS, D) QoL. BDI – Beck depression inventory, CES-D – Center for epidemiological studies depression scale, EDS - excessive daytime sleepiness, ESS – Epworth sleepiness scale, HADS-A – hospital anxiety and depression scale – anxiety subscale, HADS-D – hospital anxiety and depression scale – depression subscale, HAM-D – Hamilton rating scale for depression, MMPI – Minnesota multiphasic personality inventory – depression subscale, MMPI-Pt – Minnesota multiphasic personality inventory- psychasthenia subscale, PAP – positive airway pressure, POMS-D – profile of mood states – depression subscale, POMS-T – profile of mood states -tension and anxiety subscale, QoL - quality of life, QSQ-DS - Quebec sleepiness questionnaire – daytime sleepiness, SDS – Zung self-rated depression scale, SF-36 - short form health survey (all subscales), WHO-5 - WHO-five well-being index.

Table 4
PAP versus oral placebo.

Outcome measure	No. of studies	No. of participants	Statistical method	Effect size	Heterogeneity (I^2)	GRADE assessment
Depression	5	219	Hedge's g (Random, 95% CI)	0.355 [0.187–0.524]	0.29%	Low
Anxiety	5	219	Hedge's g (Random, 95% CI)	0.225 [0.093–0.357]	0%	Low
EDS	4	187	Hedge's g (Random, 95% CI)	0.608 [0.090–1.126]	0%	Moderate
QoL	5	219	Hedge's g (Random, 95% CI)	0.491 [0.126–0.855]	33.95%	Low
AHI/RDI	1	114	Hedge's g (Random, 95% CI)	–	–	–
AI	1	114	Hedge's g (Random, 95% CI)	–	–	–

AHI – apnea hypopnea index, AI - arousal index, EDS – excessive daytime sleepiness, GRADE - grades of recommendation, assessment, development and evaluation, QoL – quality of life, RDI - respiratory disturbance index.

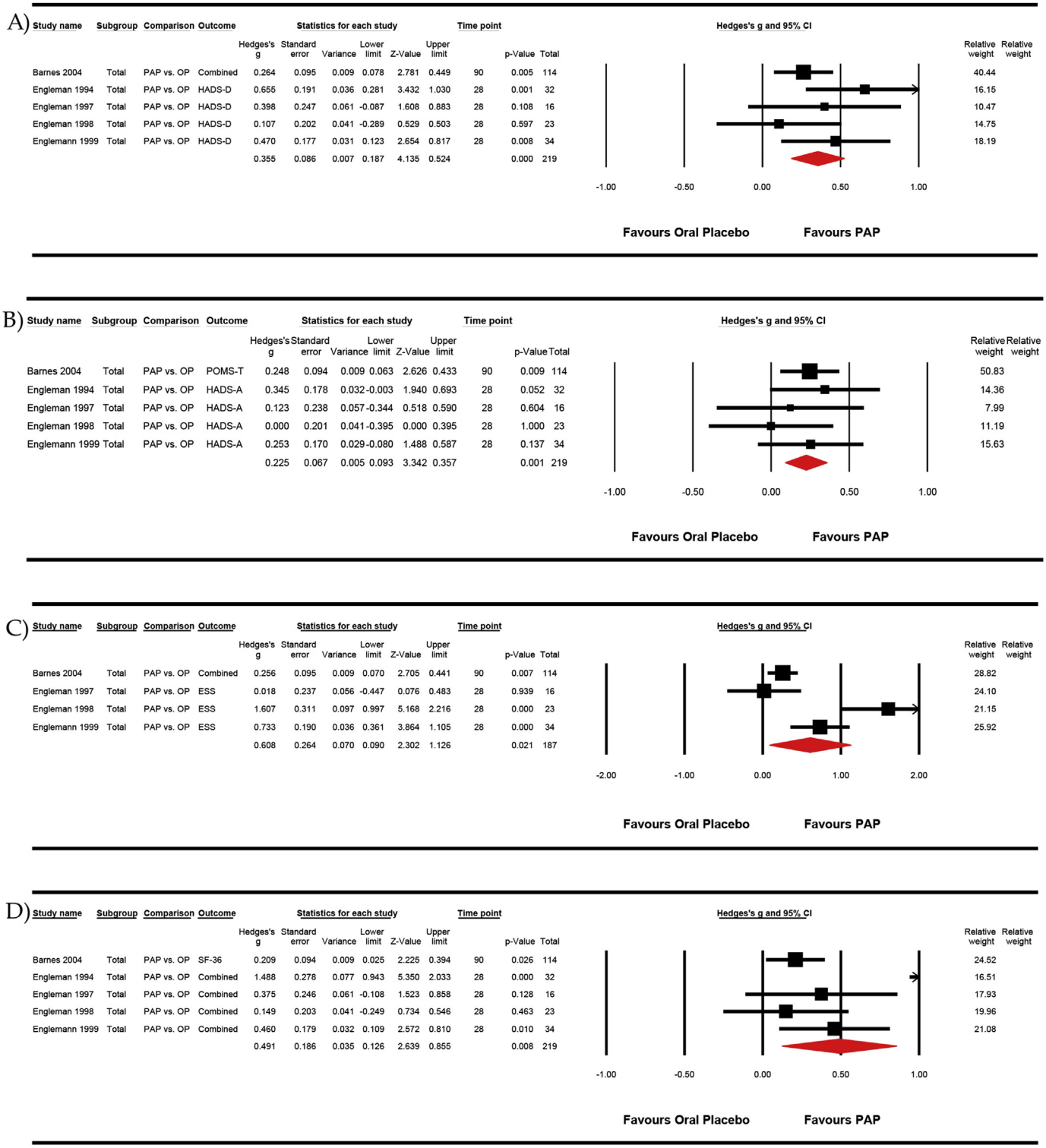


Fig. 4. PAP vs. oral placebo Forest plots. A) Depression, B) Anxiety, C) EDS, D) QoL. EDS – excessive daytime sleepiness, ESS – Epworth sleepiness scale, HADS-A – hospital anxiety and depression scale – anxiety subscale, HADS-D – hospital anxiety and depression scale – depression subscale, OP – oral placebo, PAP – positive airway pressure, POMS-T – profile of mood states -tension and anxiety subscale, QoL – quality of life, SF-36 - short form health survey (all subscales).

Table 5
PAP versus dental appliance.

Outcome measure	No. of studies	No. of participants	Statistical method	Effect size	Heterogeneity (I^2)	GRADE assessment
Depression	2	132	Hedge's g (Random, 95% CI)	0.107 [−0.72–0.287]	0%	Low
Anxiety	2	132	Hedge's g (Random, 95% CI)	0.101 [0.078–0.280]	0%	Low
EDS	2	132	Hedge's g (Random, 95% CI)	0.058 [−0.121–0.237]	0%	Low
QoL	2	132	Hedge's g (Random, 95% CI)	0.021 [−0.157–0.200]	0%	Low
AHI/RDI	2	132	Hedge's g (Random, 95% CI)	0.902 [0.692–1.113]	0%	Low
AI	2	132	Hedge's g (Random, 95% CI)	0.473 [0.285–0.661]	0%	Low
FOSQ	1	114	Hedge's g (Random, 95% CI)	–	–	–

AHI – apnea hypopnea index, AI - arousal index, FOSQ – functional outcomes of sleep questionnaire, EDS – excessive daytime sleepiness, GRADE - grades of recommendation, assessment, development and evaluation, QoL – quality of life, RDI - respiratory disturbance index.

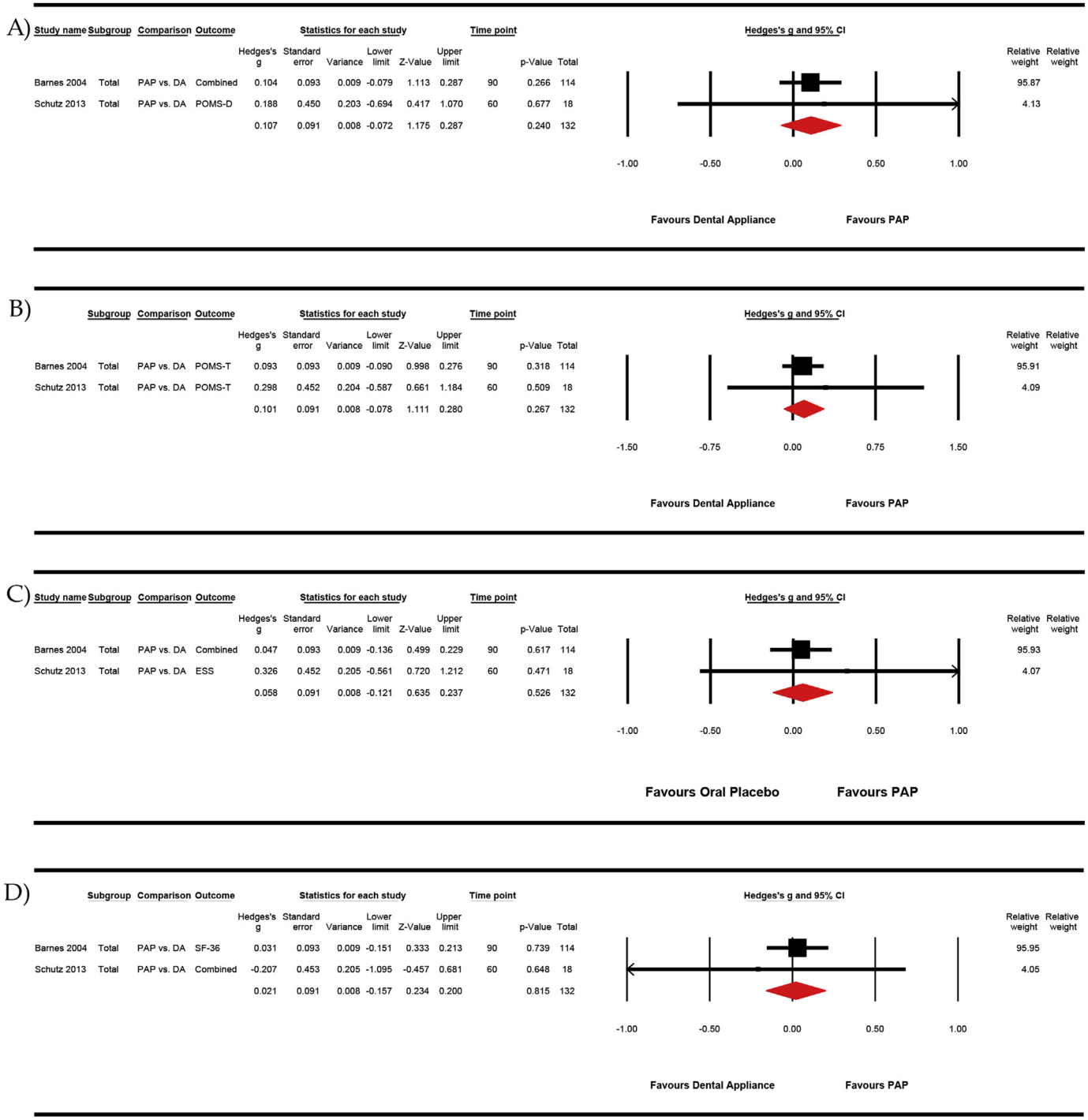


Fig. 5. PAP vs. dental appliances Forest plots. A) Depression, B) Anxiety, C) EDS, D) QoL. DA – dental appliance, EDS – excessive daytime sleepiness, ESS – Epworth sleepiness scale, PAP – positive airway pressure, POMS-D – profile of mood states – depression subscale, POMS-T – profile of mood states -tension and anxiety subscale, QoL – quality of life, SF-36 - short form health survey (all subscales).

Table 6
PAP versus sham CPAP.

Outcome measure	No. of studies	No. of participants	Statistical method	Effect size	Heterogeneity (I^2)	GRADE assessment
Depression	4	169	Hedge's g (Random, 95% CI)	−0.049 [−0.292–0.194]	0%	Low
Anxiety	4	169	Hedge's g (Random, 95% CI)	−0.073 [−0.315–0.169]	0%	Low
EDS	1	31	Hedge's g (Random, 95% CI)	–	–	–
QoL	1	31	Hedge's g (Random, 95% CI)	–	–	–
AHI/RDI	2	84	Hedge's g (Random, 95% CI)	1.881 [1.370–2.391]	0%	High
Mean SaO_2	2	39	Hedge's g (Random, 95% CI)	0.575 [−0.420–1.570]	0%	Low
AI	–	–	Hedge's g (Random, 95% CI)	–	–	–
FOSQ	1	31	Hedge's g (Random, 95% CI)	–	–	–

AHI – apnea hypopnea index, AI - arousal index, EDS - excessive daytime sleepiness, FOSQ - functional outcomes of sleep questionnaire, GRADE - grades of recommendation, assessment, development and evaluation, mean SaO_2 – mean oxygen saturation, QoL – quality of life, RDI - respiratory disturbance index.

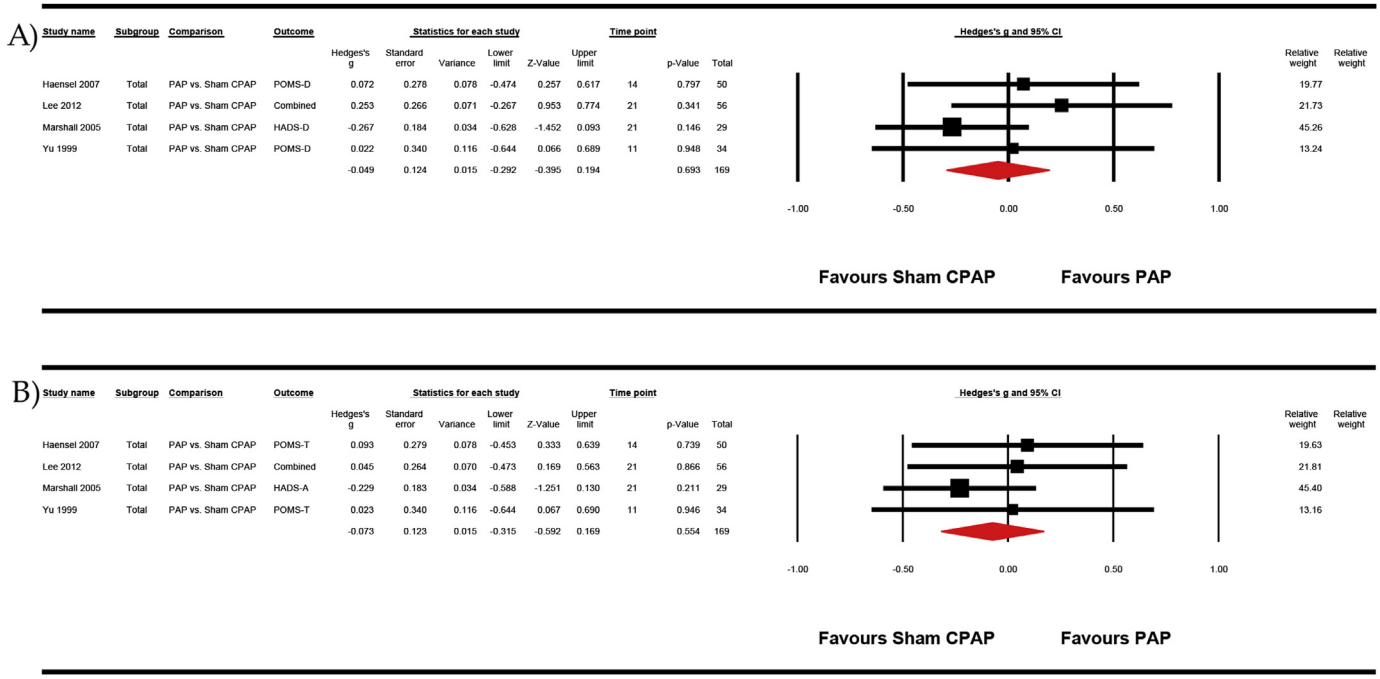


Fig. 6. PAP vs. sham CPAP Forest plots. A) Depression, B) Anxiety. CPAP – continuous positive airway pressure, HADS-A – hospital anxiety and depression scale – anxiety subscale, HADS-D – hospital anxiety and depression scale – depression subscale, PAP – positive airway pressure, POMS-D – profile of mood states – depression subscale, POMS-T – profile of mood states -tension and anxiety subscale.

therefore subject to different types of biases than the measurement of AHI by PSG. The decision to only include validated psychometric scales in this analysis, without commonly accepted substitutes such as the SF-36 mental health subscale, should hopefully also help to combat this issue. This meta-analysis was also conducted using individual comparisons between each placebo and active comparator control and PAP. This means that oral placebo and sham PAP were assessed independently, rather than as pooled placebo groups, where another meta-analysis has distinguished between parallel and cross-over study methodology [19]. The participants in the oral placebo trials were participants in crossover trials where the distinction between PAP and the oral treatment received would be apparent to the participants, whereas the participants in the sham PAP trials were more effectively blinded to the trial arms.

This meta-analysis was also unable to assess the impact of severe baseline depression versus subclinical depression. The studies used by Povitz et al. to assess severe baseline depressive symptoms were excluded from our analysis as one used the SF-36 mental health subscale and the other was conducted in stroke patients [19]. It would have been advantageous to compare the effect in severe and mild depression, as these populations are likely to have different subjective responses to PAP.

Conclusions

The effect of PAP on symptoms of depression and anxiety, EDS and QoL in patients with OSA was moderate. PAP was superior to oral placebo, but not to sham PAP in RCTs for depression and anxiety. PAP was also equivalent to dental appliances as a therapeutic option for improving subjective symptoms of depression and anxiety. PAP had a significant effect on the improvement of EDS which moderately favored PAP over oral placebo, but was equivalent to dental appliances. PAP was superior to all comparators for improvement in AHI, AI, and mean SaO_2. The relationship between improvement in subjective symptoms of OSA and PAP is complicated by the low overall quality of the evidence. Future randomized controlled trials between PAP and sham devices or dental appliances will be needed to determine if the effects of PAP are specific to improvements in respiratory variables, or subjective improvements due to the placebo effect of interaction with healthcare providers.

Practice points

- PAP improves depressive symptoms in patients with OSA, but it is not superior to dental appliances or sham PAP.
- The beneficial effect of PAP may be greater in patients with higher depression severity scores.
- If symptoms of depression do not improve after PAP therapy with improvement of AHI, the depressive symptoms will need to be treated using standard psychiatric therapies while the patient's OSA symptoms are being monitored.

Research agenda

- Randomized controlled trials should be conducted comparing the efficacy of PAP and sham PAP or active therapy in participants with clinical depression versus patients with subthreshold severity scores to determine the relative efficacy of PAP in these populations.
- Outcome measures such as depression, anxiety, EDS and quality of life should be included in all future trials of PAP in comparison to active and sham control arms for OSA.
- Study of the correlation between baseline AHI and response of subjective symptoms of depression and anxiety to PAP may help to clarify the link between disease severity and subjective patient outcomes.
- Future RCTs should consider evaluating treatment by time effects to determine the significance of length of treatment in the improvement in subjective symptoms of OSA such as depression.

Conflicts of interest

The authors have no conflict of interest to declare.

Funding sources/sponsors

None.

Acknowledgements

None.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.smrv.2015.07.002.

References

[1] The international classification of sleep disorders. Diagnosis and coding manual. 2nd ed. Westchester, Illinois: American Academy of Sleep Medicine; 2005.
[2] (ICSD-3) International classification of sleep disorders. 3rd ed. Westchester, Illinois: American Academy of Sleep Medicine; 2014.
[3] Epstein LJ, Kristo D, Strollo Jr PJ, Friedman N, Malhotra A, Patil SP, et al. Clinical guideline for the evaluation, management and long-term care of obstructive sleep apnea in adults. J Clin Sleep Med 2009;5:263–76.
[4] Kushida CA, Littner MR, Hirshkowitz M, Morgenthaler TI, Alessi CA, Bailey D, et al. Practice parameters for the use of continuous and bilevel positive airway pressure devices to treat adult patients with sleep-related breathing disorders. Sleep 2006;29:375–80.
[5] Gay P, Weaver T, Loube D, Iber C. Evaluation of positive airway pressure treatment for sleep related breathing disorders in adults. Sleep 2006;29:381–401.
[6] Giles TL, Lasserson TJ, Smith BH, White J, Wright J, Cates CJ. Continuous positive airways pressure for obstructive sleep apnoea in adults. Cochrane Database Syst Rev 2006. CD001106.
[7] Guilleminault C, Van der Hoed J, Mitler M. Clinical overview of the sleep apnea syndrome. In: Guilleminault C, Dement WC, editors. Sleep apnea syndromes. New York: AR Liss; 1978. p. 1–12.
*[8] Gupta MA, Simpson FC. Obstructive sleep apnea and psychiatric disorders: a systematic review. J Clin Sleep Med 2015;11:165–75.
[9] Lopresti AL, Drummond PD. Obesity and psychiatric disorders: commonalities in dysregulated biological pathways and their implications for treatment. Prog Neuropsychopharmacol Biol Psychiatry 2013;45:92–9.
[10] Penninx BW, Milaneschi Y, Lamers F, Vogelzangs N. Understanding the somatic consequences of depression: biological mechanisms and the role of depression symptom profile. BMC Med 2013;11:129.
*[11] Saunamaki T, Jehkonen M. Depression and anxiety in obstructive sleep apnea syndrome: a review. Acta Neurol Scand 2007;116:277–88.
[12] Andrews JG, Oei TP. The roles of depression and anxiety in the understanding and treatment of obstructive sleep apnea syndrome. Clin Psychol Rev 2004;24:1031–49.
[13] Kapur VK, Baldwin CM, Resnick HE, Gottlieb DJ, Nieto FJ. Sleepiness in patients with moderate to severe sleep-disordered breathing. Sleep 2005;28:472–7.
[14] Mediano O, Barcelo A, de la Pena M, Gozal D, Agusti A, Barbe F. Daytime sleepiness and polysomnographic variables in sleep apnoea patients. Eur Respir J 2007;30:110–3.
[15] Roure N, Gomez S, Mediano O, Duran J, Pena Mde L, Capote F, et al. Daytime sleepiness and polysomnography in obstructive sleep apnea patients. Sleep Med 2008;9:727–31.
[16] Bjornsdottir E, Keenan BT, Eysteinsdottir B, Arnardottir ES, Janson C, Gislason T, et al. Quality of life among untreated sleep apnea patients compared with the general population and changes after treatment with positive airway pressure. J Sleep Res 2015;24:328–38.
*[17] Luyster FS, Buysse DJ, Strollo Jr PJ. Comorbid insomnia and obstructive sleep apnea: challenges for clinical practice and research. J Clin Sleep Med 2010;6:196–204.
*[18] Sanchez AI, Martinez P, Miro E, Bardwell WA, Buela-Casal G. CPAP and behavioral therapies in patients with obstructive sleep apnea: effects on daytime sleepiness, mood, and cognitive function. Sleep Med Rev 2009;13:223–33.
*[19] Povitz M, Bolo CE, Heitman SJ, Tsai WH, Wang J, James MT. Effect of treatment of obstructive sleep apnea on depressive symptoms: systematic review and meta-analysis. PLoS Med 2014;11:e1001762.
[20] Gupta M, Simpson F, Lyons D. The effect of treating obstructive sleep apnea with positive airway pressure on symptoms of depression and anxiety: a systematic review and meta-analysis. PROSPERO 2014 [cited 3 February 2014] http://www.crd.york.ac.uk/PROSPEROFILES/7419_PROTOCOL_20140101.pdf.
[21] Moher D, Liberati A, Tetzlaff J, Altman DG, The PG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med 2009;6:e1000097.
[22] Version 5. 1. 0 Cochrane handbook for systematic reviews of interventions. The Cochrane Collaboration; 2011.
[23] Version Review manager (RevMan) [Computer program]. Version 5.3. [computer program]. Copenhagen, Denmark: The Cochrane Collaboration; 2014.
[24] Bornstein M, Hedges LV, JPT H, Rothstein HR. Introduction to meta-analysis. United Kingdom: John Wiley & Sons, Ltd; 2009.
[25] Version Comprehensive meta-analysis [Computer program] version 2.2.064 [computer program]. Englewood, NJ: BioStat Inc.; 2011.
[26] Version 3.2 for Windows GRADEpro. [Computer program] version 3.6.1 [computer program]. Hamilton, Ontario: McMaster University; 2014.
[27] Associação Fundo de Incentivo à Pesquisa. Continuous positive airway pressure and oral appliances treatments in mild obstructive sleep apnea (OSAS). USA: National Library of Medicine; 2000 [cited 2014 June 20], http://clinicaltrials.gov/show/NCT01461486.
[28] Barnes M, McEvoy RD, Banks S, Tarquinio N, Murray CG, Vowles N, et al. Efficacy of positive airway pressure and oral appliance in mild to moderate obstructive sleep apnea. Am J Respir Crit Care Med 2004;170:656–64.
[29] Borak J, Cieslicki JK, Koziej M, Matuszewski A, Zielinski J. Effects of CPAP treatment on psychological status in patients with severe obstructive sleep apnoea. J Sleep Res 1996;5:123–7.
[30] Brigham and Women's Hospital. Apnea, bariatric surgery versus continuous positive airway pressure (CPAP) trial (ABC). USA: National Library of Medicine; 2000 [cited 2014 June 20], http://clinicaltrials.gov/show/NCT01187771.
[31] Campos-Rodriguez F. Effect of CPAP treatment in women with moderate-to-severe OSA. USA: National Library of Medicine; 2000 [cited 2014 June 20], http://clinicaltrials.gov/show/NCT02047071.
[32] Canessa N, Castronovo V, Cappa SF, Aloia MS, Marelli S, Falini A, et al. Obstructive sleep apnea: brain structural changes and neurocognitive

* The most important references are denoted by an asterisk.

function before and after treatment. Am J Respir Crit Care Med 2011;183:1419–26.
[33] Castronovo V, Canessa N, Strambi LF, Aloia MS, Consonni M, Marelli S, et al. Brain activation changes before and after PAP treatment in obstructive sleep apnea. Sleep 2009;32:1161–72.
[34] Derderian SS, Bridenbaugh RH, Rajagopal KR. Neuropsychologic symptoms in obstructive sleep apnea improve after treatment with nasal continuous positive airway pressure. Chest 1988;94:1023–7.
[35] Diamanti C, Manali E, Ginieri-Coccossis M, Vougas K, Cholidou K, Markozannes E, et al. Depression, physical activity, energy consumption, and quality of life in OSA patients before and after CPAP treatment. Sleep Breath 2013;17:1159–68.
[36] El-Sherbini AM, Bediwy AS, El-Mitwalli A. Association between obstructive sleep apnea (OSA) and depression and the effect of continuous positive airway pressure (CPAP) treatment. Neuropsychiatr Dis Treat 2011;7:715–21.
[37] Engleman HM, Kingshott RN, Wraith PK, Mackay TW, Deary IJ, Douglas NJ. Randomized placebo-controlled crossover trial of continuous positive airway pressure for mild sleep apnea/hypopnea syndrome. Am J Respir Crit Care Med 1999;159:461–7.
[38] Engleman HM, Martin SE, Deary IJ, Douglas NJ. Effect of continuous positive airway pressure treatment on daytime function in sleep apnoea/hypopnoea syndrome. Lancet 1994;343:572–5.
[39] Engleman HM, Martin SE, Deary IJ, Douglas NJ. Effect of CPAP therapy on daytime function in patients with mild sleep apnoea/hypopnoea syndrome. Thorax 1997;52:114–9.
[40] Engleman HM, Martin SE, Kingshott RN, Mackay TW, Deary IJ, Douglas NJ. Randomised placebo controlled trial of daytime function after continuous positive airway pressure (CPAP) therapy for the sleep apnoea/hypopnoea syndrome. Thorax 1998;53:341–5.
[41] Ferini-Strambi L, Baietto C, Di Gioia MR, Castaldi P, Castronovo C, Zucconi M, et al. Cognitive dysfunction in patients with obstructive sleep apnea (OSA): partial reversibility after continuous positive airway pressure (CPAP). Brain Res Bull 2003;61:87–92.
[42] Haensel A, Norman D, Natarajan L, Bardwell WA, Ancoli-Israel S, Dimsdale JE. Effect of a 2 week CPAP treatment on mood states in patients with obstructive sleep apnea: a double-blind trial. Sleep Breath 2007;11:239–44.
*[43] Herold J, Acker J, Richter K, Piehl A, Scholz F, Niklewski G, et al. Effect of NCPAP on depressive symptoms in OSA-patients. Am J Respir Crit Care Med 2011;183(1).
[44] Kawahara S, Akashiba T, Akahoshi T, Horie T. Nasal CPAP improves the quality of life and lessens the depressive symptoms in patients with obstructive sleep apnea syndrome. Int Med 2005;44:422–7.
*[45] Lee IS, Bardwell W, Ancoli-Israel S, Loredo JS, Dimsdale JE. Effect of three weeks of continuous positive airway pressure treatment on mood in patients with obstructive sleep apnoea: a randomized placebo-controlled study. Sleep Med 2012;13:161–6.
[46] Lee IS, Bardwell W, Ancoliisrael S, Loredo JS, Dimsdale JE. Effect of 3-week continuous positive airway pressure treatment on mood in patients with obstructive sleep apnea: a randomized placebo-controlled study. Psychosom Med 2011;73(3):A41.
*[47] Marshall NS, Neill AM, Campbell AJ, Sheppard DS. Randomised controlled crossover trial of humidified continuous positive airway pressure in mild obstructive sleep apnoea. Thorax 2005;60:427–32.
[48] O'Donoghue FJ, Wellard R, Rochford PD, Dawson A, Barnes M, Ruehland WR, et al. Magnetic resonance spectroscopy and neurocognitive dysfunction in obstructive sleep apnea before and after CPAP treatment. Sleep 2012;35:41–8.
[49] Paparrigopoulos TJ, Economou NT, Koutsourelakis I, Giannakopoulou H, Kritikou I, Perraki E, et al. Amelioration of affective disorders and quality of life in obstructive sleep apnea (OSA) patients following continuous positive airway pressure (CPAP) therapy. Eur Neuropsychopharmacol 2009;19:S345–6.
[50] Ramos Platon MJ, Espinar Sierra J. Changes in psychopathological symptoms in sleep apnea patients after treatment with nasal continuous positive airway pressure. Int J Neurosci 1992;62:173–95.
[51] Sanchez AI, Buela-Casal G, Bermudez MP, Casas-Maldonado F. The effects of continuous positive air pressure treatment on anxiety and depression levels in apnea patients. Psychiatry Clin Neurosci 2001;55:641–6.
*[52] Schutz TC, Cunha TC, Moura-Guimaraes T, Luz GP, Ackel-D'Elia C, Alves Eda S, et al. Comparison of the effects of continuous positive airway pressure, oral appliance and exercise training in obstructive sleep apnea syndrome. Clin Sao Paulo Braz 2013;68:1168–74.
[53] Schwartz DJ, Karatinos G. For individuals with obstructive sleep apnea, institution of CPAP therapy is associated with an amelioration of symptoms of depression which is sustained long term. J Clin Sleep Med 2007;3:631–5.
[54] Schwartz DJ, Kohler WC, Karatinos G. Symptoms of depression in individuals with obstructive sleep apnea may be amenable to treatment with continuous positive airway pressure. Chest 2005;128:1304–9.
[55] Stanford University. Comparative outcomes management with electronic data technology (COMET) study. USA: National Library of Medicine; 2000 [cited 2014 June 20], http://clinicaltrials.gov/show/NCT01461473.
[56] Uppsala University. Life-style changes in obstructive sleep apnea. USA: National Library of Medicine; 2000 [cited 2014 June 20], http://clinicaltrials.gov/show/NCT01102920.
[57] Yamamoto H, Akashiba T, Kosaka N, Ito D, Horie T. Long-term effects nasal continuous positive airway pressure on daytime sleepiness, mood and traffic accidents in patients with obstructive sleep apnoea. Respir Med 2000;94:87–90.
[58] Ye L, Pien GW, Ratcliffe SJ, Weaver TE. Gender differences in obstructive sleep apnea and treatment response to continuous positive airway pressure. J Clin Sleep Med 2009;5:512–8.
*[59] Yu BH, Ancoli-Israel S, Dimsdale JE. Effect of CPAP treatment on mood states in patients with sleep apnea. J Psychiatr Res 1999;33:427–32.
[60] Feng Y, Zhang Z, Dong ZZ. Effects of continuous positive airway pressure therapy on glycaemic control, insulin sensitivity and body mass index in patients with obstructive sleep apnoea and type 2 diabetes: a systematic review and meta-analysis. NPJ Prim Care Respir Med 2015;25:15005.
[61] Iftikhar IH, Hoyos CM, Phillips CL, Magalang UJ. Meta-analyses of the association of sleep apnea with insulin resistance, and the effects of CPAP on HOMA-IR, adiponectin, and visceral adipose fat. J Clin Sleep Med 2014;11:475–85.
[62] Crawford MR, Bartlett DJ, Coughlin SR, Phillips CL, Neill AM, Espie CA, et al. The effect of continuous positive airway pressure usage on sleepiness in obstructive sleep apnoea: real effects or expectation of benefit? Thorax 2012;67:920–4.
[63] Schulz KF, Altman DG, Moher D. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. J Clin Epidemiol 2010;63:834–40.